Available online at www.sciencedirect.com
SciVerse ScienceDirect
ELSEVIER
Procedia Technology 4 (2012) 624 – 628
Procedia Technology

C3IT-2012

LoPP: Locality Preserving Projections for Moving Object Detection

M.T. Gopala Krishna[a], V.N. Manjunath Aradhya[c], M. Ravishankar[a], D.R. Ramesh Babu[b]

[a,c] *Dept. of ISE, Dayananda Sagar College of Engineering, Bangalore, India*
[b] *Dept. of CSE, Dayananda Sagar College of Engineering, Bangalore, India*
Email: [a] *gopalmtm@gmail.com,* [c] *aradhya1980@yahoo.co.in*

Abstract

Automatic moving object detection and tracking is very important task in video surveillance applications. In this paper, we propose a novel scheme for moving object detection based on Locality Preserving Projections (LPP). It is also known as Laplacian eigenmaps, which optimally preserves the neighborhood structure of the data set [1]. The proposed method was tested on standard **PETS** dataset and many real time video sequence and the results was satisfactory.

Keywords: Object Detection; Object Tracking; Subspace Learning; LPP.

1. Introduction

The increasing need for automated video analysis has generated a great deal of interest in object detection and tracking algorithms. The three important key steps in video analysis are detection of interesting moving objects, tracking of such objects from frame to frame, and analysis of object tracks to recognize their behavior. The major potential applications includes in motion-based recognition, automated surveillance, video indexing, human-computer interaction, traffic monitoring, vehicle navigation etc. [2]. Every tracking method requires an object detection mechanism either in every frame or when the object first appears in the video. Several common object detection methods have been proposed in the literature, which are based on point detectors, segmentation, background modeling and supervised classifiers [2]. Finding interest points in images have been long used in the context of motion, stereo, and tracking problems. A comparative evaluation of interest point detectors can be seen in [3].

From the literature, one can see that background modeling methods, especially statistical methods offer more robustness to illumination changes and dynamic backgrounds. The statistical background modeling methods can be categorized into: Gaussian Models, Support Vector Models and Subspace Learning Models [4]. Methods pertains to Gaussian Models are: Single Gaussian (SG) [5], Single General Gaussian (SGG) [6], Mixture of Gaussian (MOG) [7], Mixture of General Gaussians (MOGG) [8], and Kernel Density Estimation (KDE) [9]. The second category uses more sophisticated statistical models as support vector machines (SVM) [10], support vector regression (SVR)[11], and support vector data description (SVDD) [12]. The third category utilizes Subspace Learning Methods. Subspace Learning using Principal Component Analysis was first used by Oliver et al. [13]. Many variants to PCA can also be seen in literature.

2212-0173 © 2012 Published by Elsevier Ltd. Open access under CC BY-NC-ND license.
doi:10.1016/j.protcy.2012.05.100

Other Subspace Learning Methods such as Independent Component Analysis (ICA) [14], Incremental Non-Negative Matrix Factorization [15], and Incremental Rank Tensor [16] can be seen in literature.

As an alternative to the Principal Component Analysis, the Locality Preserving Projections (LPP) also known as Laplacian eigenmaps, was proposed which optimally preserves the neighborhood structure of the data set [1]. It is a a classical linear technique that projects the data along the directions of maximal variance. As a result, LPP shares many of the data representation properties of nonlinear techniques such as Laplacian Eigenmaps or Locally Linear Embedding. Motivated by the above facts, in this work we explore the idea of LPP to detect moving objects in video. The organization of the paper is as follows: Section 2 provides the detail description about LPP. In section 3, experiment results and comparative study is analyzed. Finally, conclusion is drawn at the end.

2. Proposed Method

Unlike from Principal Component Analysis (PCA), main objective of LPP [1] is to preserve the local structure of the input vector space by explicitly considering the manifold structure. Because it preserves the neighborhood information, its classification performance is much finer than other subspace approach like PCA. The generic problem of linear dimensionality reduction is the following. Let there be N number of input data points $(d_1, d_2, \cdots, d_N)$, which are in $\Re^M$. In the first step of this algorithm is to construct the adjacency graph **G** of N nodes, such that node i and j are linked if d_i and d_j are *close* with respect to each other in any of the following two conditions:

1. k-nearest neighbors: Nodes i and j are linked by an edge, if i is among k-nearest neighbors of j or vice-versa.
2. ϵ-neighbors: Nodes i and j are linked by an edge if $\left\|d_i - d_j\right\|^2 < \epsilon$, where $\|\cdot\|$ is the usual Euclidean norm.

Next step is to construct the weight matrix **Wt**, which is a sparse symmetric $N \times N$ matrix with weights Wt_{ij} if there is an edge between nodes i and j, and 0 if there is no edge. Two alternative criterion to construct the weight matrix:

1. Heat-Kernel: $Wt_{ij} = e^{\frac{-\left\|d_i - d_j\right\|^2}{t}}$, if i and j are linked.
2. $Wt_{ij} = 1$, iff nodes i and j are linked by an edge.

The objective function of LPP model is to solve the following generalized eigenvalue-eigenvector problem:

$$XLX^T a = \lambda XDX^T a \tag{1}$$

Where **D** is the diagonal matrix with entries as $D_{ii} = \sum_j w_{ji}$ and $L = D - W$ is the laplacian matrix.

The transformation matrix **W** is formed by arranging the eigenvectors of Eq.(1) ordered according to their eigenvalues, $\lambda_1 < \lambda_2, \ldots, < \lambda_l$. Thus, the feature vector y_i of input d_i is obtained as follows:

$$d_i \rightarrow y_i = A^T d_i \;\; \forall i = 1, 2, \ldots, N \tag{2}$$

Note: The XDX^T matrix is always singular because of high-dimensional nature of the image space. To alleviate this problem, PCA is used as the preprocessing step to reduce the dimensionality of the input vector space.

The above said is generic approach, for object detection the problem can be summarized as follows:

Learning Phase:.

- Two images (Previous and Current) is acquired, the average image, μ_b, is then computed and two images mean-subtracted;
- Follow the steps described in section 2, from constructing adjacency graph to PCA.
- The covariance matrix is computed and the best K eigenvectors stored in an eigenvector matrix, ϕ_{Kb}.

Table 1. Percentage of False Alarms

Methods	% of False Alarms
PCA	0.08
Proposed Method	0.065

Classification Phase:.

- Current and Previous (CP) Image are projected onto Eigenspace as $I' = \phi_{Kb}$ - μ_b.
- Final objects are detected by projecting Eigenvectors onto I'.

3. Experiment Results & Comparative Study

To evaluate the performance of our algorithm, we use few video sequences which differ in the issue raised by the applications, such as people counting, vehicle monitoring, environmental variation from which objects are to be detected. The publically available **PETS 2000, 2001 & 2002** dataset are used for this purpose. There are a relatively few object occurrences and object interactions in the dataset. The frequency of periods of partial and full occlusion between objects is also quite low. One of the challenging features of this dataset is the small distant objects that have a very low contrast relative to the background [17]. Sample images are shown in Figure 1.

The succeeding comparative methodology compares the performance of our algorithm against the original eigen-background method [13]. In Figure 2 we can see the foreground motion detections achieved on those four video sequences. Row 2 is the eigen-background method; Row 3 shows the results obtained with the algorithm proposed in this article. Both the methods usually require a post-processing step before labeling connected components, but in this work we have not come out with any operation. We also tested our algorithm on another video sequence which is shown in figure 3. The detected objects from our method is comparatively better compared to standard PCA approach. Interesting point is on the shape of the detected regions are some how similar to the original ones. Whereas in PCA method, it fails to preserve the structure of the object (Please refer figure 3). If the shape of the objects is not preserved properly the future steps may lead to inaccurate results.

We also tested on many real time video sequence which includes brighter light conditions, terrace video sequence in clouding conditions, with noisy environment, etc,. The proposed algorithm gave satisfactory results compared to PCA [13] based method. We notice that the computation time of the proposed method is better compared to PCA[13] based method. We also noticed that false alarms regions obtained from the PCA method is high compared to our proposed method. Detected blobs from the proposed method is quite favorable compared to PCA method. For this purpose, we calculated the percentage of false alarms on the video sequence (200 frames). Table 1 shows the percentage of false alarms achieved from both the algorithms. From the table it is clear that the percentage of false alarms obtained from the proposed algorithm is quite good compared to PCA method. The performance is due to the inherent advantages of proposed method, which optimally preserves the neighborhood structure of the data set provides a greater insight compared to PCA technique.

Fig. 1. PETS Dataset

4. Conclusion

As an alternative to the Principal Component Analysis, the Locality Preserving Projections (LPP) also known as Laplacian eigenmaps, was proposed which optimally preserves the neighborhood structure of the data set [1]. It is a classical linear technique that projects the data along the directions of maximal variance. As a result, LPP shares many of the data representation properties of nonlinear techniques such as Laplacian Eigenmaps or Locally Linear Embedding. Motivated by the above facts, in this work we explore the idea of LPP to detect moving objects in video. The proposed method was compared to with known PCA based method and showed encouraging results. Work on different issue raised by the applications and with intrinsic difficulties of the sequence will be our future work.

References

[1] X. He, S. Yan, Y. Hu, P. Niyogi., Face recognition using laplacianfaces, IEEE Trans on Pattern Analysis and Machine Intelligence 27 (3) (2005) 328–340.
[2] A. Yilmaz, O. Javed, M. Shah, Object Tracking: A Survey, ACM Computing Surveys 38 (4) (2006) 1–45.
[3] K. Mikolajczyk, C. Schmid, A performance evaluation of local descriptors, in: In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2003, pp. 1615–1630.
[4] T. Bouwmans, Subspace learning for background modeling: A survey, Recent Patents on Computer Science 2 (3) (2009) 223–234.
[5] C. Wren, A. Azarbayejani, T. Darrell, A. Pentland, Pfinder: Real-time tracking of the human body, IEEE Trans on Pattern Analysis and Machine Intelligence 19 (7) (1997) 780–785.
[6] H. Kim, R. Sakamoto, I. Kitahara, T. Toriyama, K. Kogure, Robust silhouette extraction technique using background subtraction, in: 10th Meeting on Image Recognition and Understand (MIRU), 2007.
[7] C. Stauffer, W. Grimson, Adaptive background mixture models for real-time tracking, in: IEEE Conf on CVPR, 1999, pp. 246–252.
[8] M. Allili, N. Bouguila, D. Ziou, A robust video foreground segmentation by using generalized gaussian mixture modelling, in: Fourth Canadian Conf on Computer and Robot Vision (CRV), 2007, pp. 503–509.
[9] A. Elgammal, D. H. D, L. Davis, Non-parametric model for background subtraction, in: 6th European Conf on Computer Vision, 2000, pp. 751–767.
[10] H. Lin, T. Liu, J. Chuang, A probabilistic svm approach for background scene initialization, in: International Conference on Image Processing, 2002, pp. 893–896.
[11] J. Wang, G. Bebis, R. Miller, Robust video-based surveillance by integrating target detection with tracking, in: IEEE Workshop on Object Tracking and Classification Beyond the Visible Spectrum in conjunction with CVPR, 2006, pp. 137–142.
[12] A. Tavakkoli, M. Nicolescu, G. Bebis, A novelty detection approach for foreground region detection in videos with quasi-stationary backgrounds, in: ISVC, 2006, pp. 40–49.
[13] N. M. Oliver, B. Rosario, A. Pentland, A bayesian computer vision system for modeling human interactions, IEEE Trans on Pattern Analysis and Machine Intelligence 22 (8) (2000) 831–843.
[14] M. Yamazaki, G. Xu, Y. Chen, Detection of moving objects by independent component analysis, in: ACCV, 2006, pp. 467–478.
[15] S. Bucak, B. Gunsel, O. Gursoy, Incremental non-negative matrix factorization for dynamic background modelling, in: International Workshop on Pattern Recognition in Information System, 2007.
[16] X. Li, W. Hu, Z. Zhang, X. Zhang, Robust foreground segmentation based on two effective background models, in: MIR, 2008, pp. 223–228.
[17] http://www.cvg.rdg.ac.uk/slides/pets.html.

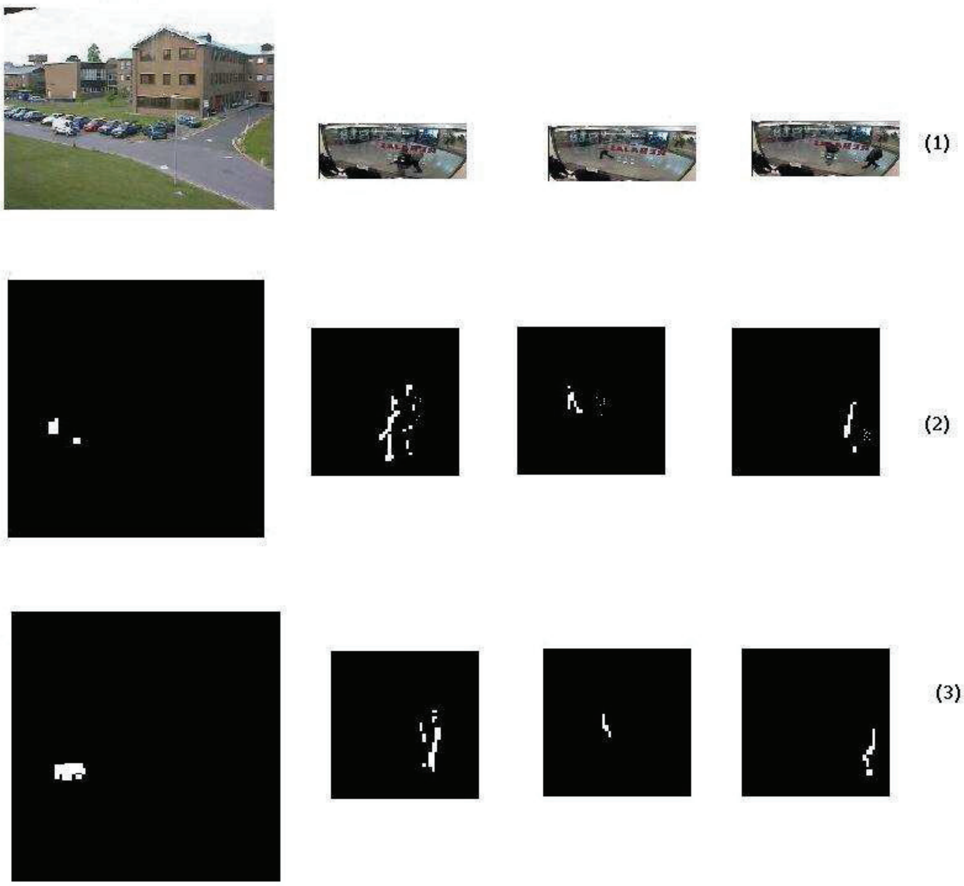

Fig. 2. Results obtained on 4 sequences with 2 algorithms: (1) Input sequence, (2) PCA Method (3) Our Algorithm

Fig. 3. Results obtained on another set of video sequence with 2 algorithms: (1) Input sequence, (2) PCA Method (3) Our Algorithm